

SEQ ID NO:1. Human Hepsin DNA Sequence: The GenBank Accession No. for human hepsin mRNA (cDNA) is M18930:

5
1 TCGAGCCCGC TTTCCAGGGA CCCTACCTGA GGGCCACAG GTGAGGCAGC CTGGCCTAGC
61 AGGCCCCACG CCACCGCCTC TGCCTCCAGG CCGCCCGCTG CTGCGGGGCC ACCATGCTCC
121 TGCCAGGCC TGGAGACTGA CCCGACCCCG GCACTACCTC GAGGCTCCGC CCCACCTGC
181 TGGACCCAG GGTCCCACCC TGGCCAGGA GGTGAGCCAG GGAATCATTG ACAAGAGGCA
241 GTGACATGGC GCAGAAGGAG GGTGGCCGGA CTGTGCCATG CTGCTCCAGA CCAAGGTGG
301 CAGCTCTCAC TGCAGGGGAC CTGCTACTTC TGACAGCCAT CGGGGCGGCA TCCTGGGCCA
10 361 TTGTGGCTGT TCTCCTCAGG AGTGACCAGG AGCCGCTGTA CCCAGTGCAG GTCAGCTCTG
421 CGGACGCTCG GCTCATGGTC TTTGACAAGA CGGAAGGGAC GTGGCGGCTG CTGTGCTCCT
481 CGCGCTCCAA CGCCAGGGTA GCCGGACTCA GCTGCGAGGA GATGGGCTTC CTCAGGGCAC
541 TGACCCACTC CGAGCTGGAC GTGCGAACGG CGGGCGCCAA TGGCACGTCG GGCTTCTTCT
601 GTGTGGACGA GGGGAGGCTG CCCCACACCC AGAGGCTGCT GGAGGTCATC TCCGTGTGTG
15 661 ATTGCCCCAG AGGCCGTTTC TTGGCCGCCA TCTGCCAAGA CTGTGGCCCG AGGAAGTGC
721 CCGTGGACCG CATCGTGGGA GGCCGGGACA CCAGCTTGGG CCGGTGGCCG TGGCAAGTCA
781 GCCTTCGCTA TGATGGAGCA CACCTCTGTG GGGGATCCCT GCTCTCCGGG GACTGGGTGC
841 TGACAGCCGC CCACTGCTTC CCGGAGCGGA ACCGGGTCTT GTCCCGATGG CGAGTGTGTTG
901 CCGGTGCCGT GGCCAGGCC TCTCCACAG GTCTGCAGCT GGGGGTGCAG GCTGTGGTCT
20 961 ACCACGGGGG CTATCTTCCC TTTCCGGACC CCAACAGCGA GGAGAACAGC AACGATATTG
1021 CCCTGGTCCA CCTCTCCAGT CCCCTGCCCC TCACAGAATA CATCCAGCCT GTGTGCCTCC
1081 CAGTGCCCGG CCAGGCCCTG GTGGATGGCA AGATCTGTAC CGTGACGGGC TGGGGCAACA
1141 CGCAGTACTA TGGCCAACAG GCCGGGGTAC TCCAGGAGGC TCGAGTCCCC ATAATCAGCA
1201 ATGATGTCTG CAATGGCGCT GACTTCTATG GAAACCAGAT CAAGCCCAAG ATGTTCTGTG
25 1261 CTGGCTACCC CGAGGGTGGC ATTGATGCCT GCCAGGGCGA CAGCGGTGGT CCCTTTGTGT
1321 GTGAGGACAG CATCTCTCGG ACGCCACGTT GCGGCTGTG TGGCATTGTG AGTTGGGGCA
1381 CTGGCTGTGC CCTGGCCCGA AAGCCAGGCG TCTACACCAA AGTCAGTGAC TTCCGGGAGT
1441 GGATCTTCCA GGCCATAAAG ACTCACTCCG AAGCCAGCGG CATGGTGACC CAGCTCTGAC
1501 CGGTGGCTTC TCGCTGCGCA GCCTCCAGGG CCCGAGGTGA TCCCGGTGGT GGGATCCACG
30 1561 CTGGGCGGAG GATGGGACGT TTTTCTTCTT GGGCCCGGTC CACAGGTCCA AGGACACCCT
1621 CCCTCCAGGG TCCTCTCTTC CACAGTGGCG GGCCCACTCA GCCCCGAGAC CACCCAACCT
1681 CACCTCCTG ACCCCCATGT AAATATTGTT CTGCTGTCTG GGACTCCTGT CTAGGTGCC
1741 CTGATGATGG GATGCTCTTT AAATAATAAA GATGGTTTTG ATT

35

SEQ ID NO:2. Human Hepsin Polypeptide Sequence (417 a.a.): The protein_id number is "AAA36013.1:"

40
NH₂-
1 MAQKEGGRTV PCCSRPKVAA LTAGTLLLLT AIGAASWAIV AVLLRSDQEP LYPVQVSSAD
61 ARLMVFDKTE GTWRLLCSSR SNARVAGLSC EEMGFLRALT HSELDVRTAG ANGTSGFFCV
121 DEGRLPHTQR LLEVISVDCD PRGRFLAAIC QDCGRRKLPV DRIVGGRDTS LGRWPWQVSL
45 181 RYDGAHLCGG SLLSGDWVLT AAHCFFERNR VLSRWRVFAG AVAQASPHGL QLGVAQVYH
241 GGYLPFRDPN SEENSNDIAL VHLSSPLPLT EYIQPVCLPA AGQALVDGKI CTVTGWGNTQ
301 YYGQQAGVLQ EARVPIISND VCNGADFYGN QIKPKMFCAG YPEGGIDACQ GDSGGPFVCE
361 DSISRTPRWR LCGIVSWGTV CALAQKPGVY TKVSDFREWI FQAIKTHSEA SGMVTQL
-COOH